The Thesis Committee for Randolph Westley Smith
Certifies that this is the approved version of the following thesis:

Short-term and Long-term Recall in Adults Who Stutter

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Courtney T. Byrd

Elizabeth H. Hampton

Short-term and Long-term Recall in Adults Who Stutter

by

Randolph Westley Smith, B.A.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2016

Acknowledgements

I would like to thank Dr. Courtney Byrd for her tremendous support, motivation, encouragement, and reassurance throughout this project. I would also like to thank Megann McGill and Zoi Gkalitsiou for their patience and assistance in the planning, implementation, and interpretation of the research involved, without which I would have been lost many times over. Thank you to Elizabeth Hampton for serving on my committee, and thank you to the participants who were so generous with their time for making this study possible.

Abstract

Short-term and Long-term Recall in Adults Who Stutter

Randolph Westley Smith, M.A.

The University of Texas at Austin, 2016

Supervisor: Courtney T. Byrd

Research has demonstrated that working memory of adults who stutter may be compromised compared to adults who do not stutter. However, these data are largely limited to production tasks. Additionally, explorations of recall accuracy have been limited to short term recall. The purpose of the present study was to further investigate the integrity of short-term memory of adults who stutter as well as long-term memory. Participants listened to lists of words that were maximally phonologically related, maximally semantically related, and lists that were both semantically and phonologically related. In the short-term condition, lists were presented three successive times and participants were required to recall words immediately after each presentation in the form of nonvocal typed responses. In the long-term condition, lists were presented once and participants were required to recall words 20 minutes later in the form of nonvocal typed responses.

Table of Contents

List of Tables

List of Figures

INTRODUCTION

Stuttering is widely understood to be a multifactorial disorder. Among the factors that contribute to this disorder, there is evidence to suggest both working memory (e.g., Baddeley, 1966a, 1966b; Bosshardt, 1990, 1993) and phonological encoding (e.g., Anderson & Byrd, 2008; Bajaj, 2007; Byrd, Vallely, Anderson, & Sussman, 2012) as contributors. Baddeley's (2003) model for working memory provides a theoretical framework for the recall accuracy of phonological information, as well as how efficiently such information is coded into both short-term memory (Baddeley, 1966b; Bosshardt, 1990, 1993) and long-term memory (Baddeley, 1966a; Moore, 1986). Reduced speed and accuracy of recall of phonological information in adults who stutter (AWS) has been observed when participants recall and identify stimuli both vocally (Byrd, Sheng, Ratner, & Gkalitsiou, 2015) and nonvocally (Byrd, McGill, & Usler, 2015), indicating possible deficits in phonological working memory as opposed to difficulties restricted to production. Given the relationship between short-term memory (STM) and long-term memory (LTM), it is reasonable to assume that persons who stutter may also demonstrate deficits in LTM. The purpose of the present study is to advance our understanding of the potential factors that contribute to stuttering by comparing short-term and long-term nonvocal recall of phonologically and semantically related word lists in adults who stutter, as well as the effect of position in a word list on short- and long-term recall accuracy.

Long-term, short-term, and working memory

Baddeley (2003) describes LTM and STM as distinct cognitive processes that comprise a unified two-component system. LTM involves neurological changes in the brain and is a more durable form of memory, while STM is a temporary storage system that functions as an "antechamber" (p. 190) to the more permanent LTM. Working memory is closely related to STM, but differs slightly in that it involves the manipulation as well as temporary storage of information. Working memory can be further defined as a multi-component system governed by the central executive, which consists of (1) the visuospatial sketchpad, (2) the episodic buffer, and (3) the phonological loop (Baddeley, 2003). The visuospatial sketchpad and the episodic buffer are beyond the scope of this paper and will not be discussed further, as they are concerned with the temporary storage and manipulation of visual information and the binding together of disparate sources of information into "chunks" or episodes, respectively. The present study explores the phonological loop as it allows for the temporary storage and manipulation of verbal and acoustic information. Additionally, the present study also investigates the central executive given that it facilitates the transfer of information between STM and LTM.

The phonological loop and its mechanisms have been described in detail by many (e.g., Baddeley, 2003; Bajaj, 2007). This aspect of working memory critically involves both a temporary phonological store and a subvocal rehearsal system that serves to maintain information within the store (Baddeley, 2003). The phonological store preserves auditory input to be remembered in the form of a phonological code, but this memory trace is susceptible to deterioration after a matter of seconds (approximately two seconds). The subvocal rehearsal system helps to preserve that trace for a longer period of time through a process of silent verbal repetition. Additionally, subvocal rehearsal serves to register nameable visual information (e.g., letters or numbers) for immediate recall, enabling visually presented items to be phonologically coded and integrated into the phonological store.

Bajaj (2007) identifies some common tasks that rely on the phonological loop (e.g., forward and backward digit spans, reading span, sequencing of letters and numbers) and highlights three variables that affect the efficiency of the phonological loop and are relevant to the current study. Given the limited capacity of this temporary storage system, word length has a direct effect on the number of words that can be stored; since shorter words can be subvocally rehearsed more frequently to prevent deterioration, higher recall accuracy is typically observed for shorter words than for longer words. The phonological similarity of words also impacts recall accuracy, with phonologically similar words being more challenging to recall than phonologically dissimilar words. Finally, articulatory suppression (i.e., requiring individuals to repeat some unrelated word, e.g., “the”, while being asked to recall an unrelated target word) also results in decreased recall accuracy, presumably because subvocal rehearsal of the target word is prevented (see also Larsen & Baddeley, 2003).

While STM and its various processes and subprocesses have been described in considerable detail, the exact relation between mechanisms connecting STM and the more durable LTM remains unclear and is the topic of considerable debate. Baddeley’s (2003) perspective is that the central executive facilitates communication across these mechanisms. Described by Baddeley as the "most important but least understood component of working memory" (Bajaj, 2007; p. 221), the central executive is responsible for attentional control of this system. It manages cognition through attentional and regulation processes to interface with its three subcomponents (visuospatial sketchpad, episodic buffer, phonological loop) and with LTM to facilitate transfer of information from the STM. Given the significance of controlling and regulating attention in this process, inhibitory control – the ability to suppress an inappropriate response in uncertain or novel conditions – likely also plays a critical role in working memory as it contributes to an individual's ability to attend to given stimuli. Finally, although it will not be explored here, it is worth noting that the central executive and the various other components of working memory are crucial not only to language but to a wide array of executive processes and other cognitive functions.

Memory and recall in adults who stutter

Numerous studies have suggested increased difficulty in adults who stutter with phonological encoding (e.g., Anderson & Byrd, 2008; Byrd, Sheng, Ratner, & Gkalitsiou, 2015), phonological working memory (Bajaj, 2007, for review; Byrd Vallely, Anderson, & Sussman, 2012; Byrd, McGill, & Usler, 2015), and attentional processes and inhibitory control (Anderson & Wagovich, 2010; Eggers et al., 2013). Given the importance of phonological working memory (i.e., the phonological loop) and attentional processes (i.e., the central executive) to integrating new information into STM and LTM – at least according to Baddeley's (2003) model – it follows that STM and LTM may be similarly affected in AWS. Relying on participants' veridical recall of information can serve as an index of whether the information was successfully encoded into STM and potentially LTM, depending on the length of time between presentation of the stimulus and the recall task.

Phonological encoding is described by Levelt (1989) as the process of retrieving the sound segments in words prior to planning and executing a motor program, and significant evidence implicates this ability as a contributing factor to difficulties producing fluent speech in AWS (e.g., Anderson & Byrd, 2008; Bosshardt, 1993; Byrd, Vallely, Anderson, & Sussman, 2012). Not surprisingly, then, phonological working memory also appears to be impaired in this population (Bajaj, 2007). Bosshardt investigated phonological working memory in relation to stuttering by comparing silent and oral reading rates (1990) and silent reading rates and recall and recognition of nonwords (1993) in AWS. These studies revealed that AWS exhibited longer silent reading times, suggesting impaired subvocal rehearsal (a crucial element of phonological working memory), and those longer silent reading times were correlated with poorer performance on recall and recognition of nonwords. Collectively, these results point to deficiencies in phonological working memory in AWS.

If, as the above studies indicate, AWS experience delayed or inefficient phonological encoding of acoustic information, then the subsequent refreshing of that information (which begins to decay after two seconds) via subvocal rehearsal will also be inhibited and the target information will be inefficiently coded into STM. As such, AWS would be expected to have greater difficulty with short-term recall of phonologically related word lists relative to AWNS. Since recalling a list of semantically related items relies less heavily on phonological encoding, however, both AWS and AWNS should perform more similarly on short-term recall of semantically related lists.

While the relationship between stuttering and working memory and STM has received some attention in recent years (Bajaj, 2007, Byrd, Vallely, Anderson, & Sussman, 2012; Byrd, McGill, & Usler, 2015; Byrd, Sheng, Ratner, & Gkalitsiou, 2015), long-term recall in AWS remains relatively unstudied. Baddeley's (2003) model would suggest that impaired phonological working memory should negatively impact the encoding of information into STM and its subsequent transfer to LTM via the central executive, resulting in poorer long-term recall of phonological information relative to AWNS. Baddeley (1966a) had participants learn one of four lists of 10 words that were phonologically or semantically related (or a control list), perform an intervening task involving recall of digits for 20 minutes, and then recall as much of the learned list as possible; in different conditions, participants were allowed to learn the lists without distractions over four trials, or were required to complete intervening tasks intended to occupy STM (immediate recall of 8 digits) between learning trials. Results indicated that although phonological lists were learned more slowly than others (presumably due to the phonological similarity effect on working memory; Bajaj, 2007), they were recalled with greater accuracy 20 minutes later (suggesting that STM and LTM rely on different coding mechanisms); further, semantic lists were recalled with poorer accuracy than other lists at retest but not during learning trials, pointing to LTM impairment by semantically similar information.

Moore (1986) explored the relationship between stuttering and LTM using electroencephalography to measure hemispheric alpha asymmetries – a ratio of average hemispheric participation – in AWS and AWNS during several experimental conditions. Participants were presented with either recorded lists of individual words (without explicit phonological or semantic relatedness) or recordings of connected speech and were asked to either recall or recognize items from the stimulus type with which they were presented. The study showed that, relative to AWNS, AWS performed worse on both recognition and especially recall tasks and were more dependent on right hemisphere processing strategies. Moore (1986) posits that the right hemisphere is less capable of verbal recall than it is of recognition, and suggests this reflects differences between STM and LTM coding and retrieval between the hemispheres. This study does not directly test differences between STM and LTM in AWS, but is suggestive of a neurological basis in the differing mechanisms involved in short-term and long-term recall in this population. Effects of phonological or semantic relatedness on LTM in AWS remains an open question, and the present study aims to investigate this issue.

Since phonological working memory appears to be uniquely compromised in AWS (Bajaj, 2007), tasks requiring subvocal rehearsal and subsequent recall of phonologically related items should be significantly more difficult for this population than AWNS. Indeed, Byrd, Sheng, Ratner, and Gkalitsiou (2015) found through a Deese–Roediger–McDermott (DRM) paradigm with AWS and AWNS that recall of verbatim phonological information was deficient in AWS. Additionally, they found that list position affected recall accuracy similarly in both groups, with lowest recall of items from the middle of the list. However, since this study required participants to recall list items verbally, it is conceivable that potential motor deficits in AWS could interfere with the ability to vocally produce words for recall. For example, if an AWS experiences a stutter at the moment of producing a word for recall, the time consumed by that disfluency could impede recall in a way that does not occur in AWNS. Byrd, McGill, and Usler (2015) found that AWS were significantly less accurate than AWNS when asked to vocally recall nonwords rather than recall them nonvocally. Vocal versus nonvocal recall of semantically and phonologically related word lists has not yet been studied, but it seems likely that type of recall should affect accuracy when the target items rely primarily on phonological working memory for retention (i.e., as in a phonologically related word list).

The present study employs the DRM paradigm, which has been used extensively to study phonological working memory in other populations but has only been employed once with persons who stutter (Byrd, Sheng, Ratner, & Gkalitsiou, 2015). The DRM technique involves lists of words that are all associated in some manner with a single word not on the list. This unpresented, but related word is termed the "critical lure" (e.g., the critical lure *wet* might have a phonological list with items such as *vet*, *yet*, and *wit*, or a semantic list with items such as *damp*, *humid*, etc.). Additional hybrid lists containing half semantically and half phonologically related items are used as a control. The Fuzzy Trace Theory (FFT) holds that recall of semantically related lists should rely more on LTM (i.e., where the meanings of known words are stored) and be more accurate than the recall of phonologically related lists, which would rely primarily on working memory and the phonological loop and be subject to rapid deterioration of the memory trace. Another factor to consider is the effect of a word's position in a list on recall accuracy. Recall is generally stronger for words at the beginning and end of a list than for those in the middle.

Purpose and hypotheses

The purpose of the present study is to further investigate the integrity of short-term versus long-term memory in adults who stutter through the use of the DRM task. We pose the following three questions: (1) Does the short-term veridical recall of adults who stutter differ from adults who do not stutter? (2) Does the position of the word in the list impact veridical recall for adults who stutter in the same manner as adults who do not stutter? (3) Does the long-term veridical recall of adults who stutter differ from adults who do not stutter? If phonological working memory is impaired in AWS, then we expect that group to exhibit lower short-term recall accuracy relative to AWNS for all list types. Due to the more rigorous demands on phonological working memory involved in recalling phonologically similar information and the unique challenges with that process experienced by AWS, we further expect that both AWS and AWNS will accurately recall more words from the semantically related list than the phonological list but that AWS will have significantly lower short-term veridical recall for the phonological list as compared to AWNS. For the list position effect, we expect that both groups will exhibit similar veridical recall for words from the end of the lists (due to the recency effect) but that AWS will recall fewer words from the beginning and middle of the lists than AWNS across all list types (due to purported deficits in subvocal recall in AWS and the greater involvement of that process in recalling words from these list positions). Finally, with regards to long-term veridical recall, we expect both groups to perform similarly across all list types due to the limited involvement of phonological working memory in accessing long-term memory.

METHODS

Participants

INCLUSION CRITERIA

Participants included in the study met the following criteria: (a) native English speaker; (b) between the ages of 18 and 30 years; (c) no past or present speech or language disorders, except for stuttering in the AWS participants; (d) pass hearing screening per American-Speech-Language-Hearing Association (ASHA) guidelines; (e) pass visual acuity test per ASHA guidelines; and (f) no neurological, social, emotional, or psychiatric disturbances. All participants completed an extensive case history in which they were asked whether they had received prior diagnoses and/or treatment for any speech and/or language disorder other than stuttering, with specific questions related to potential history of articulation and/or phonological deficits. Participants were also screened for any prior or present global or specific memory difficulties, including recall of individual words during conversation, and were asked to report any past neurological, social, emotional, or psychiatric disturbances. Any medications currently being taken were also noted. No participant in either group reported any past and/or present language, neurological, emotional, or psychiatric diagnoses or treatment, nor did anyone report use of medication that would influence performance in the present study. Aside from AWS participants who had previously received therapy for stuttering, no participant reported past history of diagnosis and/or therapy or any indication of difficulty with articulation, phonology, and/or intelligibility. One AWS participant reported a history of concussions and occupational therapy services for related memory deficits, but this participant's performance on experimental tasks did not vary significantly from other participants and he was deemed eligible to be included in the present study.

In addition to meeting the criteria stated above, participants also had to perform within normal limits on the Peabody Picture Vocabulary Test—Fourth Edition (PPVT-4; Dunn & Dunn, 2007) Expressive Vocabulary Test—Second Edition (EVT-2; Williams, 2007), Comprehensive Test of Phonological Processing—Second Edition (CTOPP-2; Wagner, Torgesen, Rashotte, & Pearson, 2013) Blending Words, Nonword Repetition, and Phoneme Elision subtests, and the National Institutes of Health Toolbox for Assessment of Neurological and Behavioral Function including Receptive Vocabulary, Flanker, Pattern Comprehension, Working Memory, Dimensional Change, and Oral Reading subtests (National Institutes of Health, 2015). Administration of these formal measures ensured (a) that there were no participants in either talker group who had receptive or expressive vocabulary or cognitive skills that were below normal limits and (b) that there was a similar distribution of linguistic performance between the two talker groups being compared. Participants also completed baseline reaction time tasks to visual and auditory stimuli as well as the Purdue Pegboard Test to ensure group differences in task performance were not affected by individual variability in response time or manual dexterity. Results from independent samples *t*-tests demonstrated that the performances of the adults who stutter (M= 104.8; SD = 10.6) and adults who do not stutter (M=106.8; SD=4.09) did not significantly differ for receptive vocabulary [$t(8)=1.69$, $p = 0.229$]. *T*-tests also revealed that performances of the adults who stutter (M= 100.6; SD = 12.74) and adults who do not stutter (M=111.00; SD=9.38) did not significantly differ for expressive vocabulary [$t(8)=1.68$, $p = 0.231$]. Likewise, no significant differences were found between the talker groups for CTOPP subtests Phoneme Elision (PE), Blending Words (BW), or Rapid Digit Naming (RDN). (CTOPP- PE: AWS M=10.20, SD = 2.17; AWNS M=10.4, SD=1.34; $t(8)= 1.93$,p=.202; CTOPP- BW: AWS M=11, SD = 3.53; AWNS M=13, SD=1.87; $t(8)= .57$,p=.47; CTOPP- NWR: AWS M=10, SD = 2.35; AWNS M=10.6 SD=288; $t(8)=.247$,p=.632). Furthermore, no significant differences were observed between talker groups for NIH Vocabulary (AWS M=108.9, SD = 14.16; AWNS M=117.6, SD=8.11; $t(8)= 1.5$,p=.256), NIH Flanker (AWS M=103.52, SD = 12.09; AWNS M=105.94, SD=; $t(8)= 1.48$,p=.258), NIH Working Memory (AWS M=89.4, SD = 18.21; AWNS M=110.8, SD=8.79; $t(8)= 1.51$,p=.255), NIH Pattern Comparison (AWS M=76.98, SD = 14.18; AWNS M=80.66, SD=18.8650; $t(8)= .492$,p=.503), NIH DCC (AWS M=103.34, SD = 11.89; AWNS M=108.48, SD=12.75; $t(8)= .036$,p=.854), or NIH Reading (AWS M=118.58, SD = 7.43; AWNS M=120.46, SD=14.94; $t(8)= 2.58$,p=.147).

Demographics

Ten adults who do (n = 5 men; M = 22.4 years; SD = 4.4; range = 18-28 years) and do not stutter (n = 5 men; M = 23.2 years; SD = 3.8; range = 18-28 years) who were matched according to age (±2 years), education, and handedness met these criteria and participated in the present study.

STUTTERING SEVERITY

Speech production and stuttering severity were analyzed from two 15-min video recordings of conversational speech samples taken before beginning speech-language testing and experimental tasks. Intelligibility and production of articulation and/or phonological errors were informally assessed, as well. No participant in either talker group produced articulatory or phonological errors during conversation.

Severity ratings were exclusively assigned to each of the participants who stutter by the author using the Stuttering Severity Instrument for Children and Adults—Third Edition (SSI-3; Riley, 1994). One of the 5 participants received a very mild rating, three received a mild rating, and one received a rating of mild-moderate.

All five of the participants who stutter reported receiving speech therapy for stuttering, but no other therapy history was reported for participants in either talker group. We chose not to exclude adults on the basis of treatment history for stuttering for two reasons. First, there is no reason to believe that exposure to fluency therapy would differentially affect individual performance on the task employed in the study. Second, it is not uncommon for adults who stutter to report participation in fluency therapy, especially during the school years, so inclusion of these adults contributes to the ecological validity of this study.

Approval for the completion of this study was provided by the author's university institutional review board, and informed consent was obtained for each participant. All participants were provided monetary compensation for their participation. Participant characteristics and performance on standardized test measures are summarized in Table 1.

Table 1. Participant Demographic Information

Participant	Talker Group	Age	Gender	Handedness	Education	Ethnicity	Severity	PPVT-IV	EVT-2	Non-word Repetition
1	AWS	28	Male	Right	MA student	White	Mild	101	97	11
2	AWS	26	Male	Right	MA student	White	Mild	104	112	8
3	AWS	21	Male	Right	BS student	Asian	Very Mild	123	116	12
4	AWS	18	Male	Right	HS	Hispanic	Mild-mod	101	89	12
5	AWS	19	Male	Right	HS	Asian	Mild	95	89	7
6	AWNS	25	Male	Right	JD student	White		110	121	9
7	AWNS	28	Male	Right	MA	White		109	112	14
8	AWNS	24	Male	Right	BA	White		106	110	10
9	AWNS	18	Male	Right	BS student	Asian		109	116	13
10	AWNS	21	Male	Right	BS student	White		100	96	7

Note. Peabody Picture Vocabulary Test—Fourth Edition (PPVT–IV): Edition (EVT–2): standard score (M = 100, SD = 15). Mean age for AWS = 22.4 years; mean age for AWNS = 23.2 years. AWNS = adults who do not stutter; AWS = adults who stutter; Mild-mod = mild–moderate.

Participant	Talker Group	CTOPP Elision	CTOPP Blending	CTOPP Reading	NIH-PV	NIH-FICA	NIH-LSWM	NIH-DCCS	NIH-ORR
1	AWS	11	14	14	109	117	114	105	123
2	AWS	11	5	8	127	89	97	110	129
3	AWS	13	11	15	118	114	86	118	117
4	AWS	8	12	9	93	101	86	94	113
5	AWS	8	13	7	98	96	64	89	111
6	AWNS	12	15	13	122	110	108	109	142
7	AWNS	11	12	12	126	101	125	129	129
8	AWNS	9	12	12	123	104	113	109	107
9	AWNS	11	15	11	110	117	101	100	119
10	AWNS	9	11	12	108	97	107	96	107

Note. Comprehensive Test of Phonological Processing—Second Edition (CTOPP–2): subtest scaled scores. National Institutes of Health Toolbox iPad App Version 1.1 (NIH): standard score (*M* = 100, *SD* = 50); Picture Vocabulary Test (PV), Flanker Inhibitory Control and Attention Test (FICA), List Sorting Working Memory Test (LSWM), Dimensional Change Card Sort Test (DCCS), Oral Reading Recognition Test (OCCR). AWS = adults who stutter; AWNS = Adults who do not stutter.

Stimuli

TARGET WORD LISTS

Stimuli were sourced from word lists used by Byrd et al. (2015). Six lists were used, each with 12 words. Lists were centered on two different critical lure words. For each critical lure, two lists of 12 words each were developed: 12 phonological associates for the phonological condition, and 12 semantic associates for the semantic condition. Two control lists were also developed, each consisting of 12 words with no semantic or phonological relation to either critical lure. Additionally, the order of words in each list was varied such that the initial 4 words, medial 4 words, and final 4 words differed between participants.

In total, there were two different sets of lists of semantic associates, two different sets of lists of phonological associates, and two different sets of lists of control words. Phonotactic probabilities and lexical neighbors were calculated for all lists using the BLICK software program (Hays, 2012), and *t*-tests were performed to verify that the word lists used were not significantly different from one another. Average biphone phonotactic probability for control list 1 was .021 with a range of .002 to .081, and average lexical neighbors for this list was 7.833 with a range of 0 to 25. Average biphone phonotactic probability for control list 2 was .007 with a range of <.000 to .013, and average lexical neighbors was 32.923 with a range of 25 to 42. Average biphone phonotactic probability for the FACE phonological list was .004 with a range of <.000 to .006, and average lexical neighbors was 29.583 with a range of 14 to 56. Average biphone probability for the FACE semantic list was .024 with a range of <.000 to .151, and average lexical neighbors was 9.167 with a range of 0 to 28. Average biphone phonotactic probability for the MAIL phonological list was .006 with a range of .004 to .012, and average lexical neighbors was 36.083 with a range of 26 to 50. Average biphone phonotactic probability for the MAIL semantic list was .017 with a range of .003 to .041, and average lexical neighbors was 5.583 with a range of 0 to 16.

Table 2. Stimuli Lists

FACE		MAIL		CONTROL	
Phonological	Semantic	Phonological	Semantic	1	2
Fake	Mouth	Meal	Stamp	Paper	Calm
Vase	Expression	Nail	Deliver	Umbrella	Bite
Fuss	Nose	Mate	Receive	Tire	Duck
Faith	Eyes	Mile	Bills	Boss	Pale
Lace	Frown	Hail	Letters	Mistake	Guy
Fail	Wrinkle	Make	Send	Bottom	Fair
Fain	Makeup	Mall	Fax	Paint	Hill
Ace	Cheek	Sail	Express	Dodge	Soak
Case	Head	Veil	Post	Bowling	Pal
Fate	Mask	Mill	Zip	Horrible	Dug
Fame	Moustache	Mole	Address	Van	Bar
Race	Beard	Maid	Envelope	Marker	Bale

RECORDING

A female, native English speaker with a standard American accent recorded the stimuli using a digital Zoom H4 recorder in a soundproof booth. The recording was segmented into individual sound files, each containing single words and presented together with corresponding words in the appropriate experimental condition. Intelligibility of the spoken words was verified by asking three research assistants blind to the purpose of the study to listen to and write down all the words from the 6 lists. All three research assistants correctly identified all 72 words (see Table 2).

Procedure

Participants were tested individually in a quiet room. Each list was presented in auditory form using the SuperLab 5.0 software package, RokIt5 external stereo speakers, a Dell external monitor, and a Dell Optiplex 9020 computer running the Windows 7 Enterprise operating system. During administration of the task, the participant remained seated in a chair facing the speakers and computer monitor, which were approximately four feet away.

Prior to the presentation of the experimental lists, participants were told that they would hear the word list three times and have 40 seconds after each presentation to type as many of the words as they could remember in any order. After recalling the words a third time, they would hear the list for a fourth and final time, perform unrelated tasks for 20 minutes, and then have a final opportunity to type as many words as they could recall. At each recall opportunity, the participants had up to 40 seconds to enter their responses into a text box on the computer monitor using an external keyboard.

The order of the sets of lists was counterbalanced across the participants, and responses were recorded in SuperLab 5.0. Time to complete all experimental tasks, including unrelated "distraction" tasks in the interim between final list presentation and long-term recall opportunity, was approximately 3 hours. Participants were given the option to complete testing in a single session or in two 1.5 hour sessions on different days; all but one participant elected to complete the tasks in two sessions. Data compiled during in-session testing was aggregated into a Microsoft Excel spreadsheet for further analysis.

RESULTS

Participant responses were analyzed using a generalized estimating equation (GEE) in order to predict the odds of performance for adults who do and do not stutter based on the values of the predictors. The GEE with normal distribution was selected as the statistical model in order to extend the generalized linear model to allow for analysis of repeated measures within subjects. For the GEE analysis of accurate number of responses, the repeated subject variable was ParticipantID and the within-subject variable was Condition (i.e., control, phonological, semantic). Predictors in this analysis were TalkerGroup (i.e., adults who stutter, adults who do not stutter) and Condition. Dependent variables were short-term memory recall 1, short-term memory recall 2, short-term memory recall 3, and long-term recall. For the GEE analysis of accuracy by list position, the repeated subject variable, the within-subject variable, and the predictors were identical to the previous analysis. Dependent variables were initial list position, medial list position, and final list position. Statistical analyses were completed using IBM SPSS Statistics Version 23.

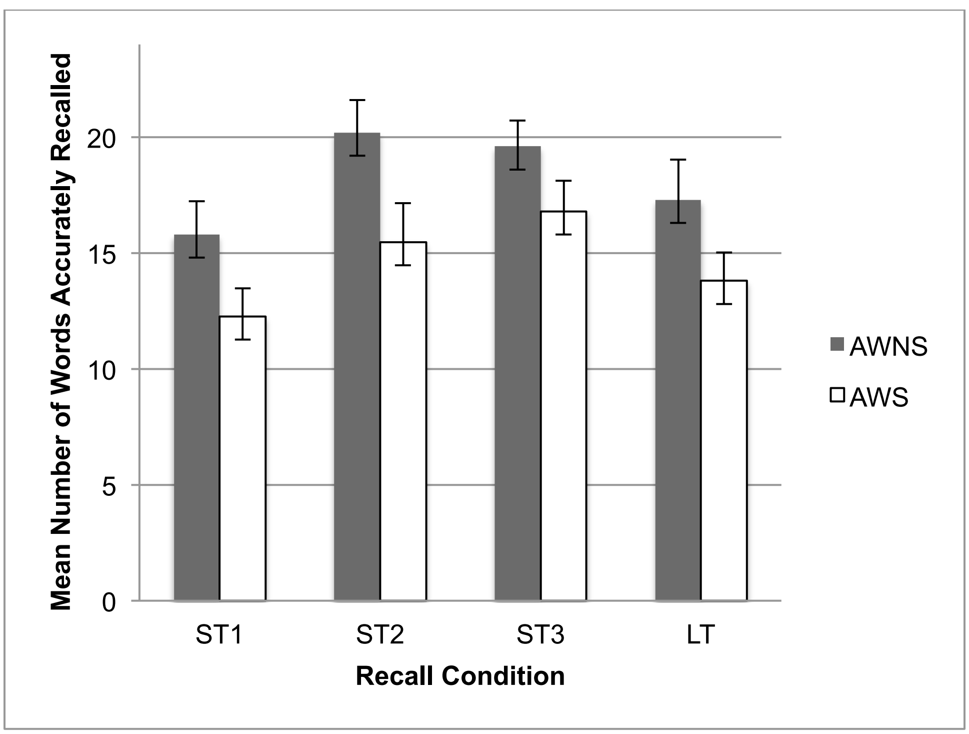
Table 3. Recall accuracy between AWS and AWNS for each condition and list position.

Dependent Variable	Wald Chi-Square	Significance	Exp(B)	95% Wald Confidence Interval for Exp(B)	
				Lower limit	Upper limit
Short-term 1	6.541	.011*	81.451	2.795	2373.331
Short-term 2	4.065	.044*	54.598	1.118	2666.287
Short-term 3	4.337	.037*	11.023	1.152	105.493
Initial list position	5.511	.019*	100	.088	1384.109
Medial list position	2.322	.128	44.701	.262	7638.353
Final list position	15.059	.000*	601.845	23.749	15251.840
Long-term	.643	.000*	3.320	.177	62.390

**indicates significance at <.05*

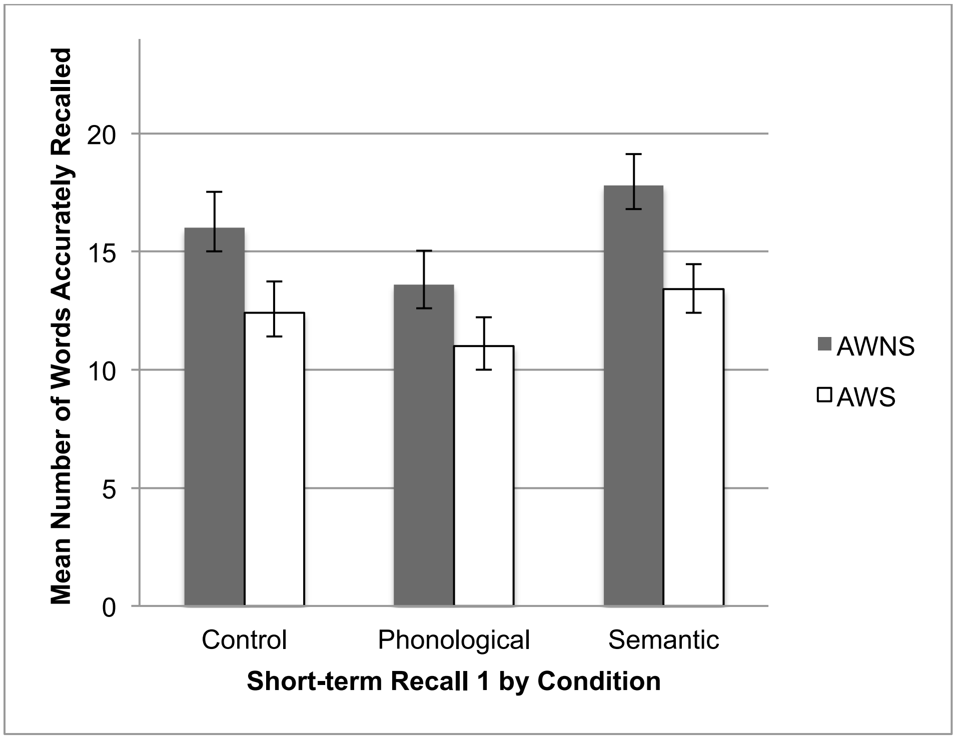
Short-term veridical recall

Figure 1. Mean and standard error for number of words accurately recalled from the three short-term recalls and one long-term recall across all list conditions for adults who stutter (AWS) versus adults who do not stutter (AWNS).



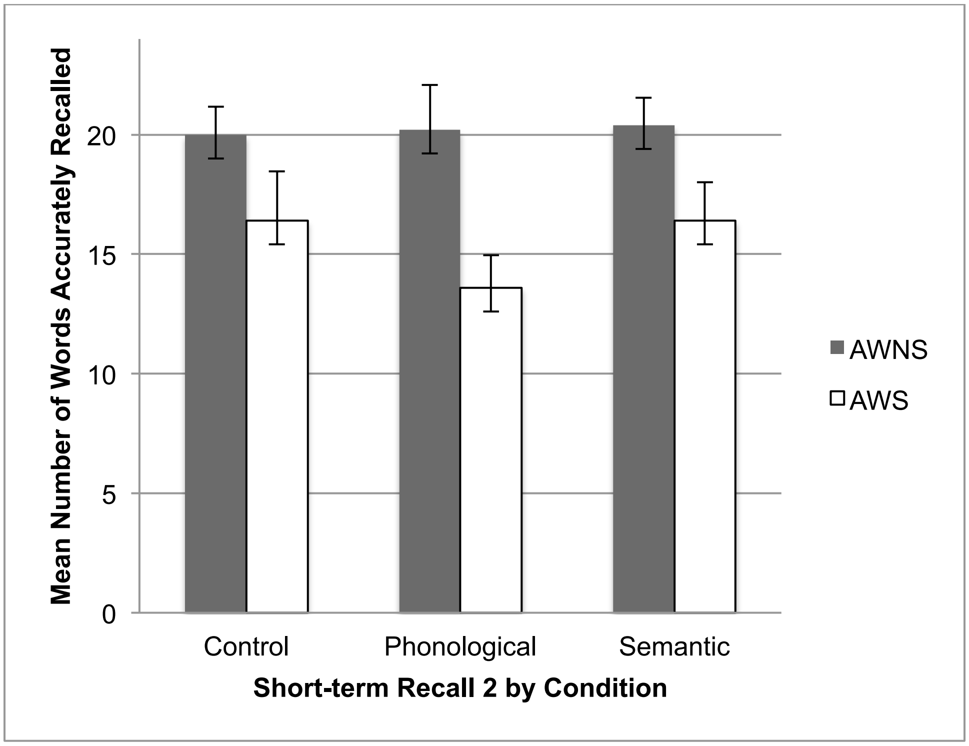
In short-term memory recall 1, adults who stutter were 81 times less likely to recall words from lists across all three conditions compared to adults who do not stutter (OR = 81.451, Wald Chi-Square = 6.541, df = 1, $p = .011$). Additionally, both adults who do and do not stutter recalled significantly more words from semantic lists (AWS: $M = 13.4$; AWNS: $M = 17.8$) than control lists (AWS: $M = 12.4$; AWNS: $M = 16.0$) or phonological lists (AWS: $M = 11.0$; AWNS: $M = 13.6$) (Wald Chi-Square = 23.632, df = 2, $p < .000$). No significant interaction between TalkerGroup and Condition was found (Wald Chi-Square = 1.739, df = 2, $p = .419$) for short-term memory recall 1. Adults who stutter were significantly less likely to accurately recall words from all list conditions compared to adults who do not stutter in the initial short-term recall condition. Additionally, all participants were significantly more likely to recall semantically related lists compared to control lists or phonologically related lists.

Figure 2. Mean and standard error bars for number of words accurately recalled from the control, phonological, and semantic list conditions during the first short-term recall for adults who stutter (AWS) versus adults who do not stutter (AWNS).



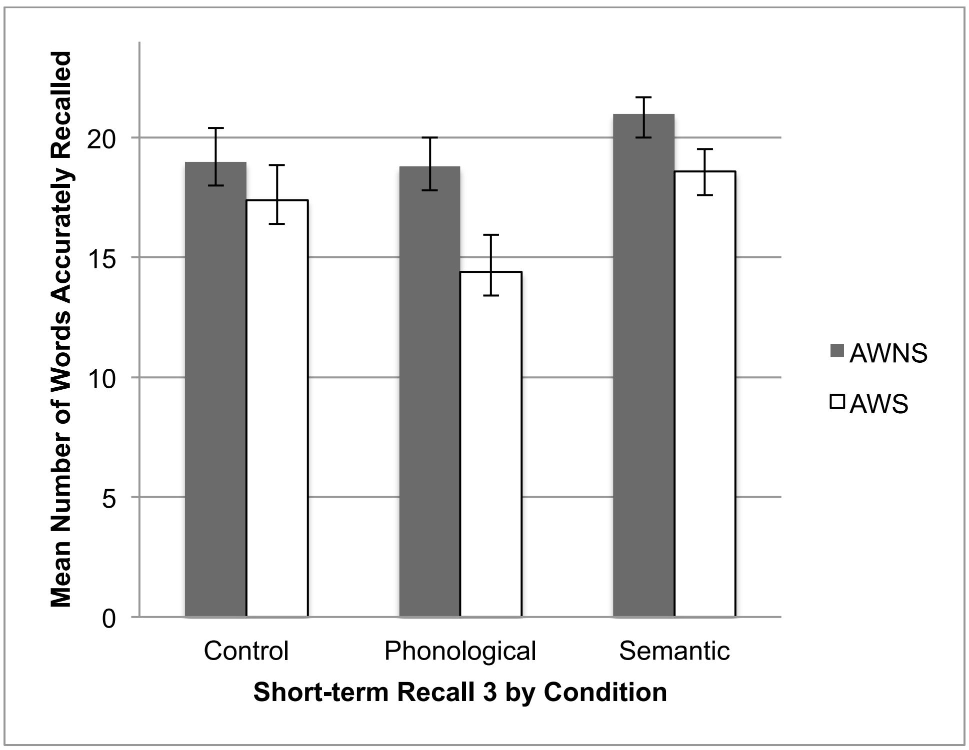
In short-term memory recall 2, adults who stutter were 54 times less likely to recall words from lists across all three conditions compared to adults who do not stutter (OR = 54.6, Wald Chi-Square = 4.065, df = 1, p = .044). No significant differences were noted for all participants across Conditions (Wald Chi-Square = 3.912, df = 2, p = .141). Additionally, no interaction between TalkerGroup and Condition was found (Wald Chi-Square = 4.774, df = 2, p = .092) for short-term memory recall 2. Thus, adults who stutter were significantly less likely to accurately recall words from all list conditions compared to adults who do not stutter in the second short-term recall task.

Figure 3. Mean and standard error bars for number of words accurately recalled from the control, phonological, and semantic list conditions during the second short-term recall for adults who stutter (AWS) versus adults who do not stutter (AWNS).



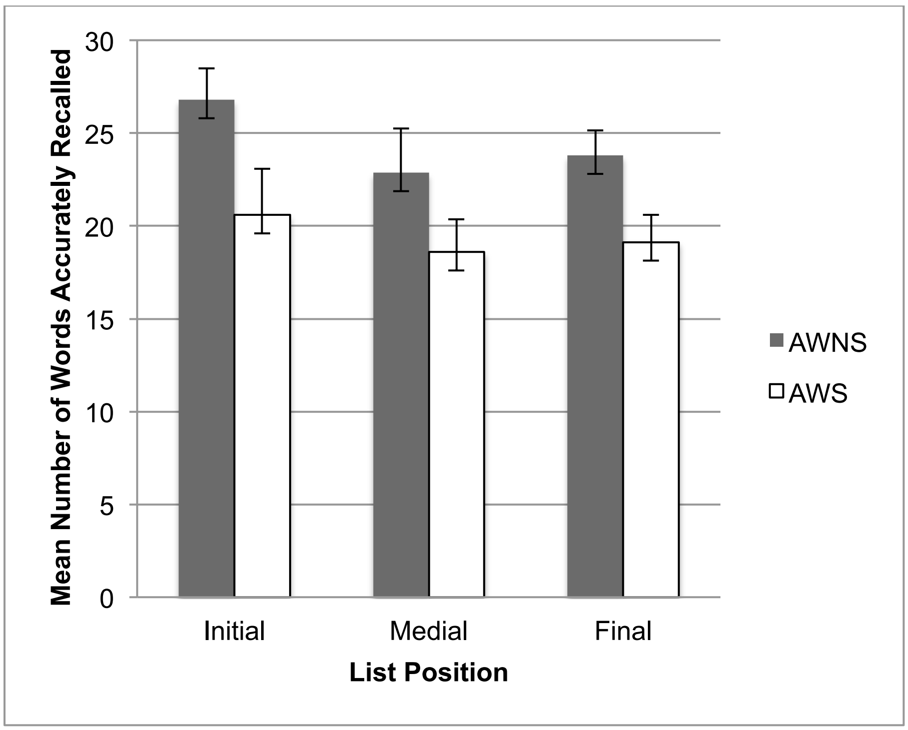
In short-term memory recall 3, adults who stutter were 11 times less likely to recall words from lists across all three conditions compared to adults who do not stutter (OR = 11.0, Wald Chi-Square = 4.337, df = 1, p = .037). Additionally, both adults who do and do not stutter recalled significantly more words from semantic lists (AWS: M = 18.6; AWNS: M = 21.0) than control lists (AWS: M = 17.4; AWNS: M = 19.0) or phonological lists (AWS: M = 14.4; AWNS: M = 18.8) (Wald Chi-Square = 16.816, df = 2, p < .000). No significant interaction between TalkerGroup and Condition was found (Wald Chi-Square = 3.201, df = 2, p = .202) for short-term memory recall 3. Thus, adults who stutter were significantly less likely to accurately recall words from all list conditions compared to adults who do not stutter in the final short-term recall condition. Additionally, all participants were significantly more likely to recall semantically related lists compared to control lists or phonologically related lists.

Figure 4. Mean and standard error bars for number of words accurately recalled from the control, phonological, and semantic list conditions during the third short-term recall for adults who stutter (AWS) versus adults who do not stutter (AWNS).



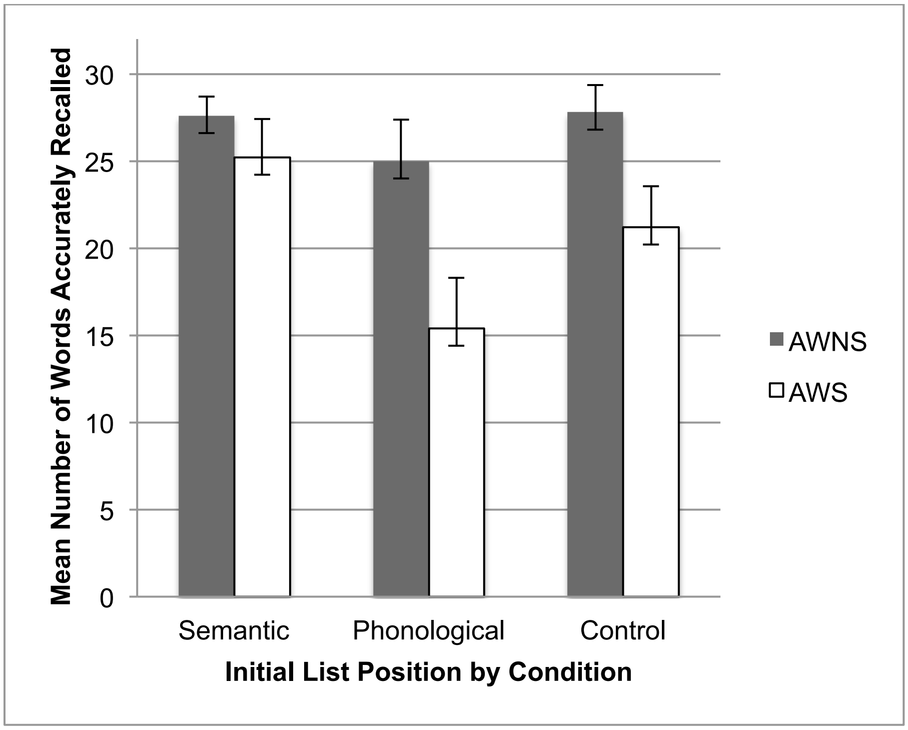
List position

Figure 5. Mean and standard error bars for number of words accurately recalled from the initial, medial, and final list positions across all conditions for adults who stutter (AWS) versus adults who do not stutter (AWNS).



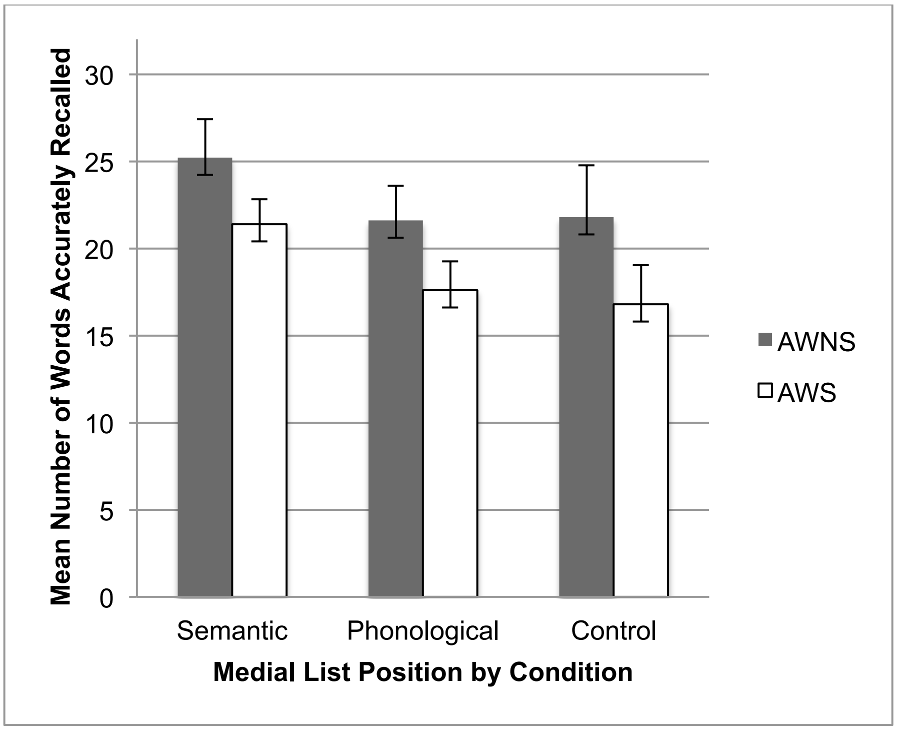
Adults who stutter were 100 times less likely to recall words in initial position (i.e., any of the first 4 words from a 12-word list) across all three conditions when compared to adults who do not stutter (OR = 100, Wald Chi-Square = 5.511, df = 1, p = .019). As a single group, all participants were 180 times less likely to recall words in initial position from the control condition compared to the semantic condition (OR = 180, Wald Chi-Square = 16.667, df = 1, p < .000). All participants were also 610 times less likely to recall words in initial position from the phonological condition compared to the control condition (OR = 610, Wald Chi-Square = 8.596, df = 1, p = .003). Additionally, adults who stutter were 66 times less likely than adults who do not stutter to recall words in initial position from the control condition compared to the semantic condition (OR = 66.686, Wald Chi-Square = 8.861, df = 1, p = .003). Adults who stutter were 1339 times less likely to recall words in initial position from the phonological condition compared to the semantic condition (OR = 1339.43, Wald Chi-Square = 4.84, df = 1, p = .028). Overall, there was a significant TalkerGroup by Condition interaction in recall of words from initial list position (Wald Chi-Square = 9.255, df = 2, p = .010). These results indicate that adults who stutter were significantly less accurate than adults who do not stutter in recalling words in initial position from all three conditions, with the least number of words recalled from the phonological condition compared to the control or semantic conditions.

Figure 6. Mean and standard error bars for number of words accurately recalled from the semantic, phonological, and control list conditions in initial list position for adults who stutter (AWS) versus adults who do not stutter (AWNS).



No significant difference was found between adults who do and do not stutter in the ability to recall words in medial position (i.e., any of the middle 4 words from a 12-word list) across all three conditions (Wald Chi-Square = 2.322, df = 1, p = .128). As a single group, all participants were 100 times less likely to recall words from medial position in the control condition compared to the semantic condition (OR = 100, Wald Chi-Square = 5.442, df = 1, p = .020). All participants were also 220 times less likely to recall words from medial position in the phonological condition compared to the semantic condition (OR = 220, Wald Chi-Square = 6.588, df = 1, p = .010). No significant interaction between TalkerGroup and Condition was found for recall of words from medial list position (Wald Chi-Square = .313, df = 2, p = .855). These findings indicate that both adults who do and do not stutter recalled words in medial position more accurately from the semantic condition than from the phonological or control conditions.

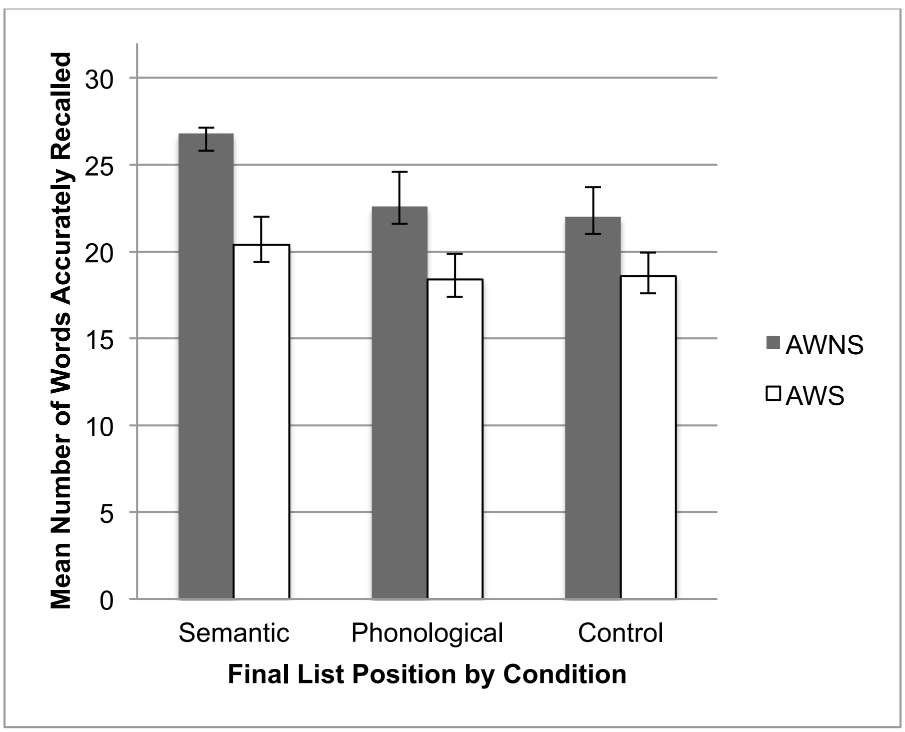
Figure 7. Mean and standard error bars for number of words accurately recalled from the semantic, phonological, and control list conditions in medial list position for adults who stutter (AWS) versus adults who do not stutter (AWNS).



Adults who stutter were 601 times less likely than adults who do not stutter to recall words from final position (i.e., any of the last 4 words from a 12-word list) across all three conditions (OR = 601.845, Wald Chi-Square = 15.059, df = 1, $p < .000$). As a single group, all participants were 121.510 times less likely to recall words in final list position from the control condition compared to the semantic condition (OR = 121.510, Wald Chi-Square = 9.796, df = 1, $p = .002$). There was no significant difference in recall accuracy at final list position between the control and phonological conditions (Wald

Chi-Square = .199, df = 1, p = .655). Additionally, no significant interaction between TalkerGroup and Condition was found for recall of words in final list position (Wald Chi-Square = 1.991, df = 2, p = .370). These results indicate that adults who stutter recalled significantly fewer words than adults who do not stutter at final list position across all three conditions. Also, both adults who do and do not stutter had significantly higher recall accuracy at final list position for the semantic condition compared to the control condition, and for the control condition compared to the phonological condition.

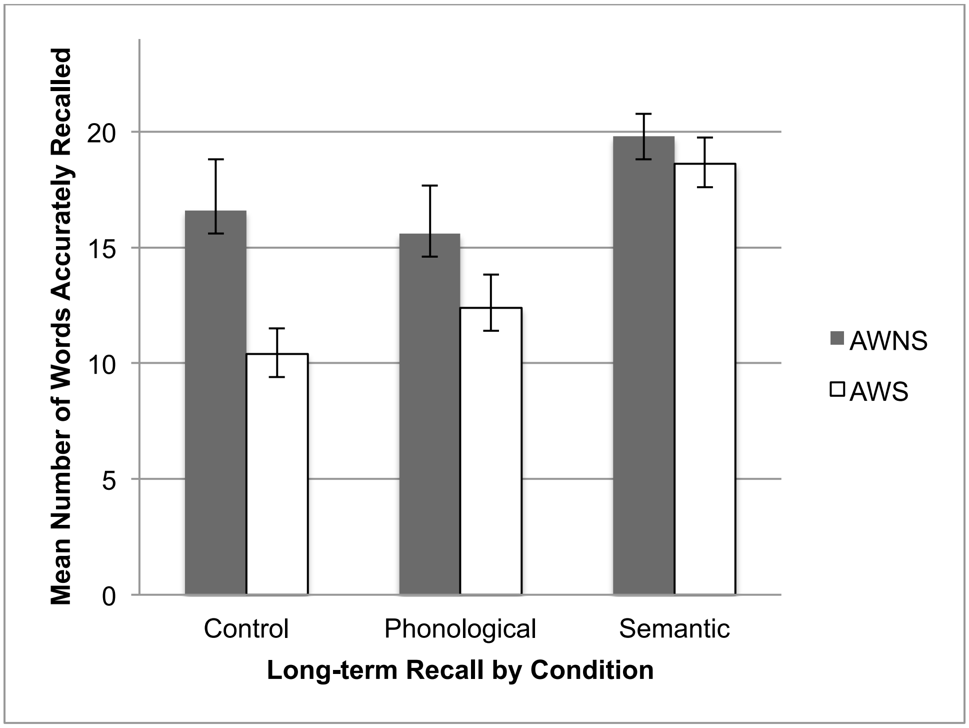
Figure 8. Mean and standard error bars for number of words accurately recalled from the semantic, phonological, and control list conditions in final list position for adults who stutter (AWS) versus adults who do not stutter (AWNS).



Long-term recall

In long-term memory recall, adults who stutter were 3 times less likely to recall words from lists across all three conditions compared to adults who do not stutter (OR = 3.320, Wald Chi-Square = .643, df = 1, p = .000). Additionally, both adults who do and do not stutter recalled significantly more words from semantic lists (AWS: M = 18.6; AWNS: M = 19.8) than control lists (AWS: M = 10.4; AWNS: M = 16.6) or phonological lists (AWS: M = 12.4; AWNS: M = 15.6) (Wald Chi-Square = 34.882, df = 2, p < .000). No significant interaction between TalkerGroup and Condition was found (Wald Chi-Square = 4.086, df = 2, p = .130) for long-term memory recall. Adults who stutter were significantly less likely to accurately recall words from all list conditions compared to adults who do not stutter in the long-term recall condition. Additionally, all participants were significantly more likely to recall semantically related lists compared to control lists or phonologically related lists.

Figure 9. Mean and standard error bars for number of words accurately recalled from the control, phonological, and semantic list conditions during the long-term recall for adults who stutter (AWS) versus adults who do not stutter (AWNS).



DISCUSSION

Research has demonstrated that phonological working memory may be uniquely compromised in persons who stutter. The present study endeavored to enhance our understanding of this deficit by investigating the integrity of short-term versus long-term memory in adults who stutter through the use of the DRM task. This paradigm was used to determine whether adults who stutter (AWS) differed from adults who do not stutter (AWNS) in their short-term recall accuracy, their long-term recall accuracy, and with regard to the list position effect on their recall accuracy. We hypothesized that AWS would exhibit lower short-term recall accuracy relative to AWNS for all list types, but that both groups would recall more words from semantically related word lists than phonologically-related or control lists. We further expected that AWS would have significantly lower short-term recall of phonologically related lists relative to AWNS. With regards to list position, we hypothesized that both AWS and AWNS would exhibit greater recall accuracy for words from the end of lists than the beginning or middle, and that AWS will recall fewer words from the beginning and middle of lists relative to AWNS across all list types. Finally, regarding long-term recall, we expected no significant differences between AWS and AWNS in recall accuracy across all list types.

Short-term veridical recall

As expected, results indicated that AWS had significantly less accurate recall compared to AWNS on semantic, phonological, and control lists on all three short-term recall opportunities. These findings are consistent with previous research (Byrd, Sheng, Ratner, & Gkalitsiou, 2015). Of note, AWS demonstrated consistent improvement in recall accuracy with each successive short-term recall opportunity, and while never achieving comparable accuracy to AWNS on a given trial, the difference in recall accuracy between the two groups was significantly smaller on the third trial compared to the first. Given demonstrated differences between AWS and AWNS in initial encoding of words via phonological working memory, AWS may have inaccurately or incompletely encoded word lists upon initial exposure but gradually improved or enhanced their working memory representations of the word lists over successive iterations.

Across all short-term recall opportunities, AWS consistently recalled the most words from the semantic condition and the fewest words from the phonological condition. We posit that this is owing to the more rigorous demands on phonological working memory involved in recalling phonologically similar information and the unique challenges with phonological encoding experienced by AWS. In contrast, semantically related word lists have a more salient organizing principle and, as operationalized in this study, are more phonologically distinct. Thus, these words compete less for encoding into working memory via the phonological loop. According to Baddeley's (2003) model of working memory, the phonological trace is subject to decay over time and will only last approximately 2 seconds. Given that the presentation of word lists in all conditions was significantly longer than 2 seconds, participants should theoretically have had to employ subvocal rehearsal to retain stimuli long enough to recall them at the end of the list presentation. This assumption suggests two possible explanations for the decreased recall accuracy of AWS on phonological lists relative to semantic and control lists. One explanation is that AWS encode the words correctly on initial exposure, but that representation diminishes substantially before the participant is provided the opportunity to recall the words a few seconds later. Alternatively, AWS may encode the words incorrectly on initial exposure and maintain that inaccurate representation via subvocal rehearsal, only to inaccurately recall those misrepresented words when given the opportunity.

Taken together, these results show that AWS do in fact exhibit lower short-term recall accuracy on all list types as a group compared to AWNS, and also that they have greater difficulty with short-term recall of phonologically related word lists relative to semantic or control lists.

List position

As described by Byrd, Sheng, Ratner, and Gkalitsiou (2015), recall accuracy is mediated by the position of words within the list. Words from the end of a list tend to be recalled most accurately, presumably because they are more accessible in short-term memory at the time of recall; this is known as the recency effect. Words from the beginning of a list are also recalled more accurately than those from the middle, presumably because there are fewer competing stimuli to recall and rehearsal strategies are more easily applied; this is known as the primacy effect. Words from the middle of a list are recalled with lower accuracy due to the absence of these two effects.

Contrary to expectations and previous research (Byrd, Sheng, Ratner, & Gkalitsiou, 2015), both AWS and AWNS recalled more words from initial position than final or medial positions across phonological, semantic, and control lists. Words from medial position were still recalled with the lowest accuracy, but the primacy effect was significantly more pronounced than the recency effect for both Talker Groups. In all list positions, AWS recalled significantly fewer words than AWNS. While the lower recall accuracy of AWS relative to AWNS is attributable to previously described phonological working memory deficits in AWS, the prominence of the primacy effect over the recency effect in both groups is an unexpected and interesting finding. One possible explanation is that participants relied heavily on rehearsal strategies and focused so intently on the first several words presented that they lacked attentional resources to attend as closely to words in medial and final list positions. Alternatively, this finding may be an anomalous byproduct of a relatively small sample size (n = 5 for each Talker Group) that could change with more robust data.

AWS recalled significantly fewer words from phonological lists relative to semantic and control lists in the initial list position. This finding is consistent with expectations and suggests that AWS struggled to encode phonologically similar stimuli relative to semantically similar or random stimuli. The precise mechanisms underlying this deficit are unclear, but may relate to rapid decay of accurate phonological representations or to maintenance of inaccurate representations through rehearsal strategies, as described earlier.

While the general trend of least recall of phonological lists and greatest recall of semantic lists was consistent across all list positions, the much lower recall of phonological lists in initial position was not observed in medial and final positions. This finding may correspond to the fact that, regardless of condition, words from medial and final positions were recalled less accurately than those in the initial position. The case may be that because participants were attending more to words in the initial position, the condition effects of phonological versus semantic recall accuracy were more apparent in this list position. If participants were unable to attend as well to medial and final list positions, it may be expected that condition effects on recall in those positions were correspondingly diminished.

Taken together, these findings show that list position mediates recall accuracy, with the primacy effect being most apparent across phonological, semantic, and control lists. Further, AWS exhibit lower recall accuracy than AWNS regardless of list position.

Long-term recall

We hypothesized that both AWS and AWNS would exhibit similar long-term recall on phonological, semantic, and control lists due to the limited involvement of phonological working memory in accessing long-term memory. Results were contrary to this expectation, with AWS exhibiting lower long-term recall relative to AWNS on all list types, and AWS recalling significantly more words from the semantic condition than from the phonological or control conditions. We suspect that the group difference in long-term recall accuracy is reflective of impaired encoding of auditory stimuli into working memory in AWS, resulting in inaccurate or inadequate transfer of information from working memory to long-term memory.

Long-term recall accuracy was highest for the semantic condition in both AWS and AWNS, and AWNS only slightly outperformed AWS for this list type. This is perhaps unsurprising, given that short-term recall accuracy for both groups was typically highest in the semantic condition as well, and that increased accuracy relative to other list types was conserved during the 20 minutes prior to participants' long-term recall opportunity. The salient organizational nature of the semantic lists was also likely a contributing factor to greater long-term recall accuracy, as participants may have been able to cue themselves to remember items from the same category more easily than items that rhymed or had no explicit relationship with one another. AWNS performed very similarly in long-term recall of control and phonological lists, suggesting that only semantic relatedness served as a helpful cue in retrieving items from long-term memory.

Interestingly, AWS had greater long-term recall of phonological lists compared to control lists. This is the only interaction for this Talker Group in which phonological lists were not the least recalled list type. One possible explanation for this unexpected finding is that, for AWS, any relationship between list items is more helpful in cueing long-term recall than no relationship at all; i.e., the fact that many words from the phonological list rhymed may have assisted in retrieval from long-term memory, whereas no such cue was available for recalling the unrelated words from the control list.

Taken together, these findings indicate long-term memory is indeed affected in AWS relative to AWNS, with semantically related information recalled at significantly greater accuracy than phonologically related information or information without an explicit relationship. Interestingly, phonologically similar information is recalled better than unrelated information, perhaps because any degree of relatedness is helpful in cueing retrieval of information from long-term memory.

Additional considerations

Although certain findings of the present study are compelling, it should be noted that the sample population used was small and lacking in diversity. Both Talker Groups consisted of 5 right-handed male participants of monolingual English background, the majority of whom had a college degree or higher. As such, any conclusions drawn here should be considered preliminary, and further research should incorporate more diverse population samples, particularly in terms of gender. That being said, some additional considerations should be made when interpreting the findings at hand.

In the present study, participants provided non-vocal typed responses during recall opportunities. Although this choice was made specifically to remove speech production from the task as a potential confound for AWS, it would be interesting to determine whether there is a difference in typed versus oral recall across list type for both short- and long-term memory performance. Another critical future consideration is how specific rehearsal strategies used by participants may affect recall accuracy across list conditions and recall type. Participants were not asked about any strategies they may have employed to help retain stimuli, and it would be interesting to assess whether specific methods resulted in different recall accuracies within groups.

CONCLUSION

The present findings suggest that phonological working memory deficits in adults who stutter may result in reduced short-term and long-term memory performance for auditory verbal stimuli. Adults who stutter exhibited reduced short-term and long-term recall of word lists relative to adults who do not stutter. Both talker groups were significantly more accurate in recalling semantically related word lists when compared to lists of words with phonologically similar forms. Additionally, both talker groups' recall of words was mediated by list position, exhibiting a pronounced primacy effect across all list conditions. Future research should explore if the short-term and long-term memory performance differ based on the modality being used to record participant responses. It would be of particular interest to compare performance across word list type when both speaker groups are required to type responses versus respond orally in a time-constrained condition.

References

Anderson, J.D., & Byrd, C.T. (2008). Phonotactic probability effects in children who stutter. *Journal of Speech, Language, and Hearing Research, 51*, 851–866.

Anderson, J.D., & Wagovich, S.A. (2010). Relationships among linguistic processing speed, phonological working memory, and attention in children who stutter. *Journal of Fluency Disorders, 35*, 216–234.

Baddeley, A.D. (1966a). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology, 18*, 302-309

Baddeley, A.D. (1966b). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18*, 362–365.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*, 189–208.

Bajaj, A. (2007). Working memory involvement in stuttering: Exploring the evidence and research implications. *Journal of Fluency Disorders, 32*, 218–238.

Bosshardt, H.G. (1990). Subvocalization and reading rate differences between stuttering and nonstuttering children and adults. *Journal of Speech and Hearing Research, 36*, 286–293.

Bosshardt, H.G. (1993). Differences between stutterers' and nonstutterers' short-term recall and recognition performance. *Journal of Speech and Hearing Research, 36,* 286–293.

Byrd, C.T., Sheng, L., Ratner, N., & Gkalitsiou, Z. (2015). Veridical and false recall in adults who stutter. *Journal of Speech, Language, and Hearing Research, 58*, 28-42.

Byrd, C.T., McGill, M., & Usler, E. (2015). Nonword repetition and phoneme elision in adults who do and do not stutter: vocal versus nonvocal performance differences. *Journal of Fluency Disorders, 44*, 17-31.

Byrd, C.T., Vallely, M., Anderson, J., & Sussman, H. (2012). Nonword repetition and phoneme elision in adults who do and do not stutter. *Journal of Fluency Disorders, 37*, 188–201.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test–Fourth Edition*. Minneapolis, MN: Pearson.

Eggers, K., De Nil, L.F. & Van Den Bergh, B.R.H. (2013). Inhibitory control in childhood stuttering. *Journal of Fluency Disorders, 38*, 1-13.

Hayes, B. (2012). BLICK – a phonotactic probability calculator (Version 1.0) [Software]. Available from http://www.linguistics.ucla.edu/people/hayes/BLICK/

Larsen, J. D., & Baddeley, A. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: Do they have a common source? *Quarterly Journal of Experimental Psychology Section A*, *56*(8), 1249-1268.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Moore, W.H. (1986). Hemispheric alpha asymmetries of stutterers and nonstutterers for the recall and recognition of words and connected reading passages: some relationships to severity of stuttering. *Journal of Fluency Disorders, 11*, 71-89.

NIH Toolbox for Assessment of Neurological and Behavioral Function (2013). Neurology, 80, Supplement 3, S1-S92. doi: 10.1212/WNL.0b013e3182872e90

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing—Second Edition*. Austin, TX: Pro-Ed.

Williams, K. T. (2007). *EVT–2: Expressive Vocabulary Test–Second Edition*. Minneapolis, MN: Pearson Assessments.